

# Аналитическое исследование: Яндекс Афиша (лето–осень 2024)

**Цель:** выявить изменения пользовательских предпочтений и популярности событий осенью 2024 года,

а также проверить гипотезы о различиях в поведении пользователей **mobile vs desktop**.

**Стек:** Python (pandas, numpy), визуализация (matplotlib / seaborn), статистика (scipy).

## Материалы проекта:

- [Ноутбук с кодом](#)
  - [README проекта](#)
  - [Репозиторий GitHub](#)
- 

## 1. Контекст и задача

Яндекс Афиша — сервис поиска мероприятий и покупки билетов. Сервис работает с партнёрами (организаторы и билетные операторы), которые размещают события и продают билеты.

### Задачи исследования:

1. провести EDA и выявить сезонные изменения (лето → осень 2024);
  2. проанализировать популярность событий, регионов и партнёров;
  3. проверить гипотезы о различиях поведения пользователей на mobile и desktop.
- 

## 2. Данные

Использовано 3 таблицы:

- **orders** — заказы билетов: user\_id, даты заказа, event\_id, device\_type, revenue, tickets\_count, days\_since\_prev и др.

- **events** — характеристики событий: тип, город, регион, площадка, организатор и др.
- **tenge\_rates** — курс kzt → rub по датам.

#### Объём:

- orders: ~290k заказов
  - events: ~22k событий
  - rates: 357 дней курса
- 

### 3. Предобработка

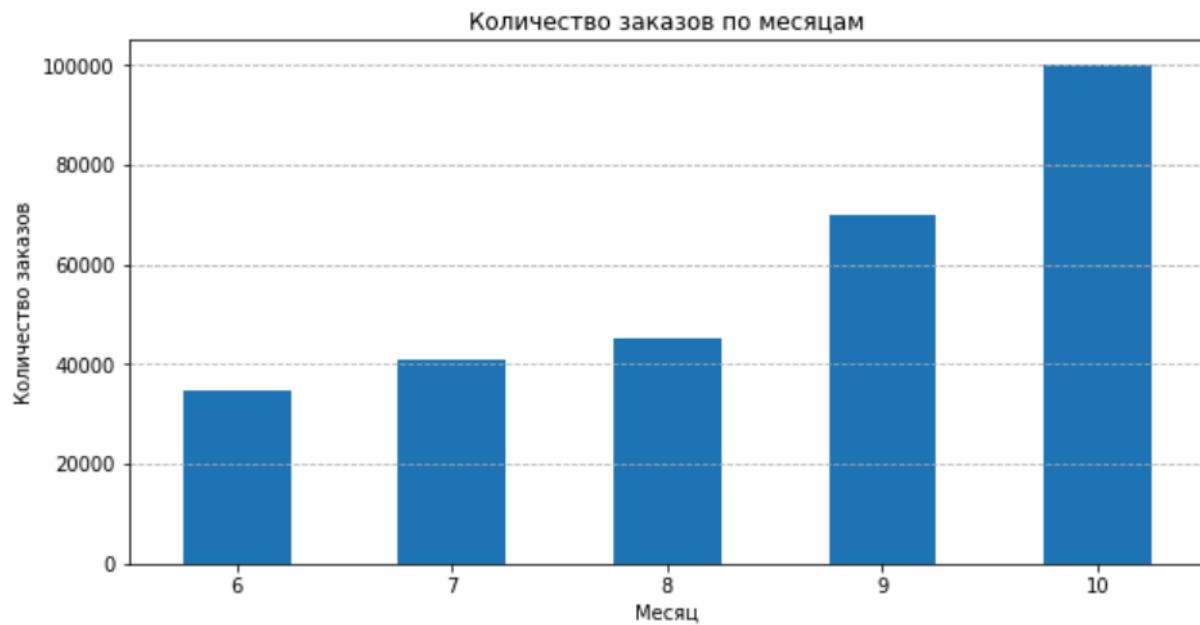
#### Ключевые шаги:

- приведение дат к datetime;
  - поиск дублей: обнаружены “неявные дубликаты” без order\_id → удалено 30 строк;
  - обработка выбросов revenue: удалены значения выше 99 перцентиля по каждой валюте (удалено 2887 заказов);
  - объединение таблиц orders + rates + events;
  - приведение выручки к одной валюте (рубли): `revenue_rub` ;
  - создание признаков:
    - `one_ticket_revenue_rub` (выручка на 1 билет),
    - `month`,
    - `season` (лето/осень/зима/весна).
- 

### 4. Исследовательский анализ (EDA)

#### 4.1 Динамика заказов по месяцам (июнь–ноябрь 2024)

Наблюдается рост числа заказов от лета к осени.

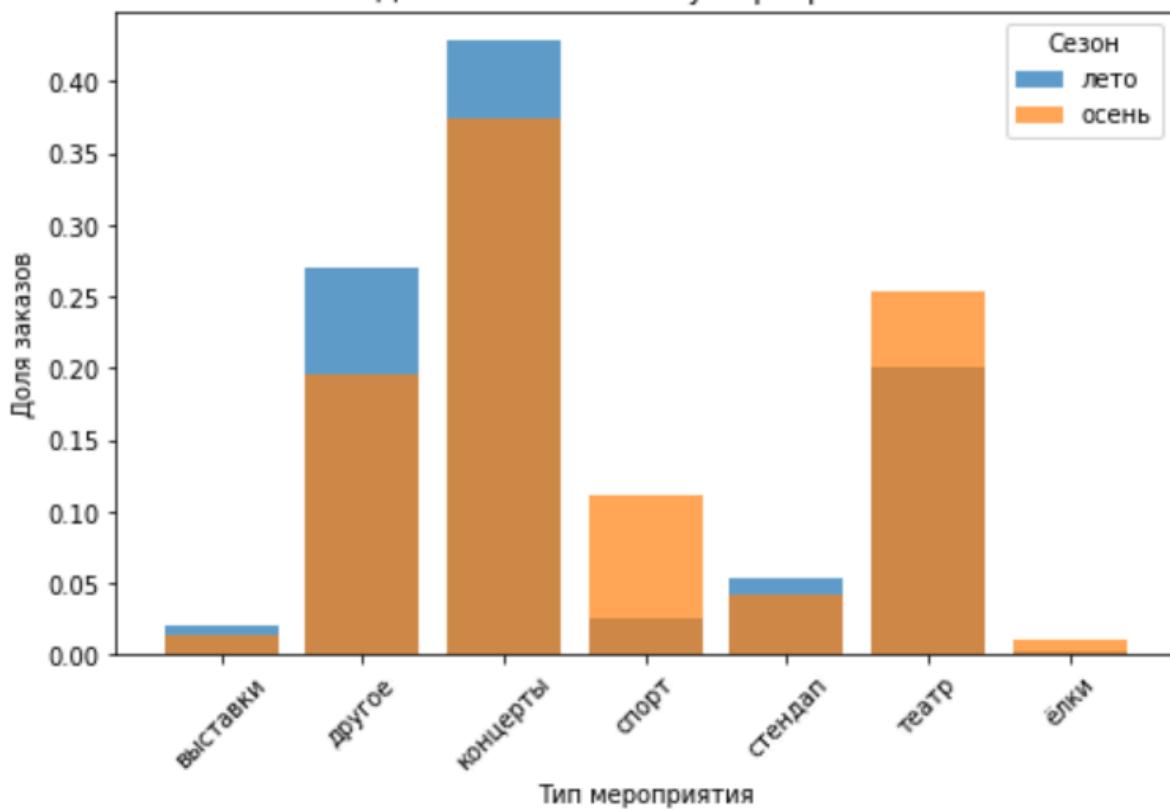


#### 4.2 Сравнение распределений летом и осенью

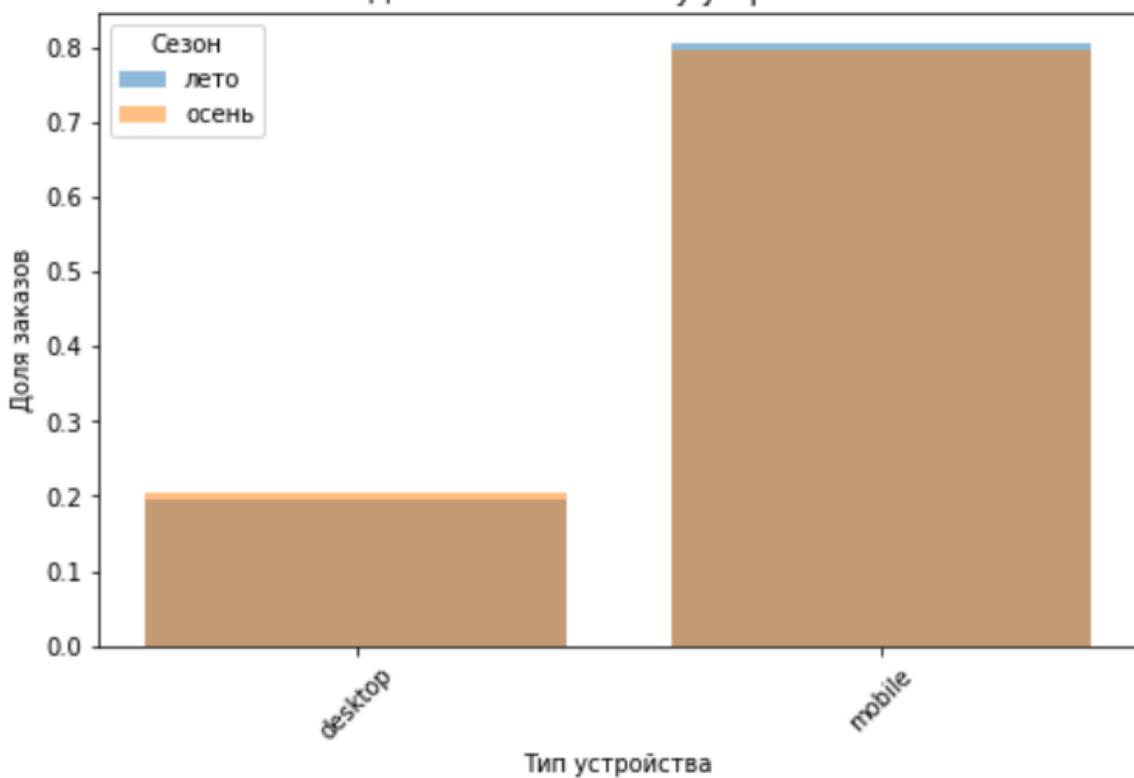
Сравнены доли заказов по:

- типам мероприятий,
- типу устройства,
- возрастному рейтингу.

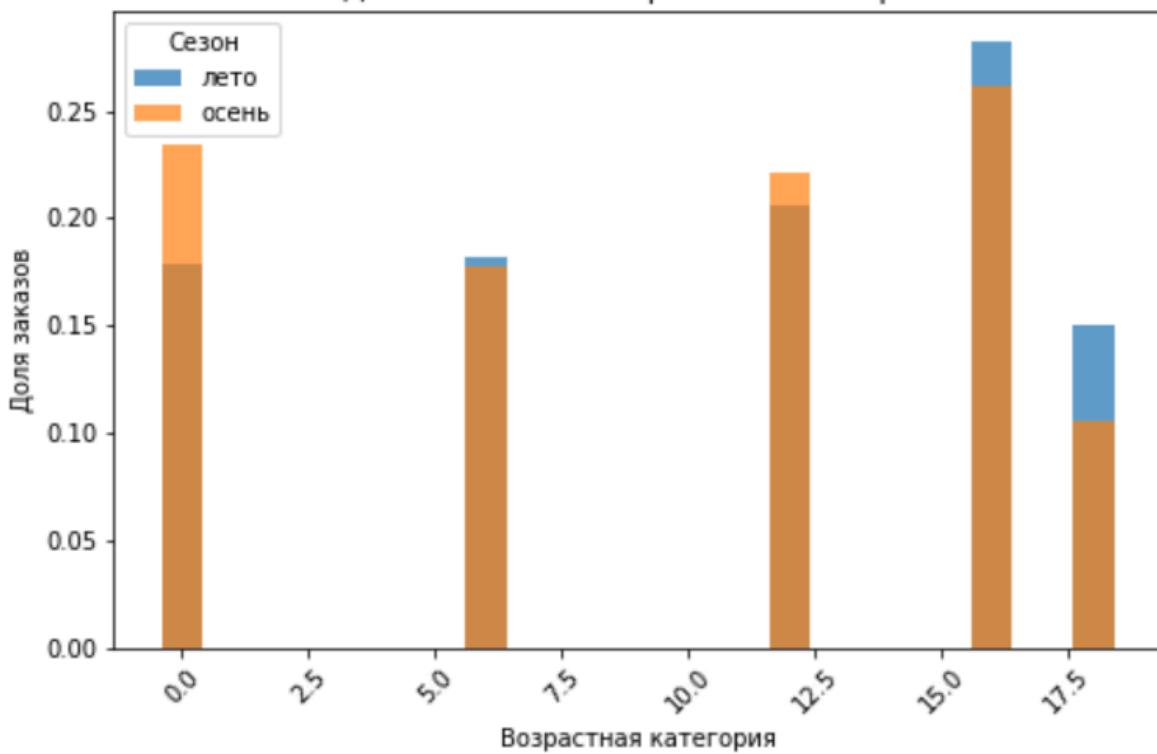
### Доли заказов по типу мероприятия



Доли заказов по типу устройства



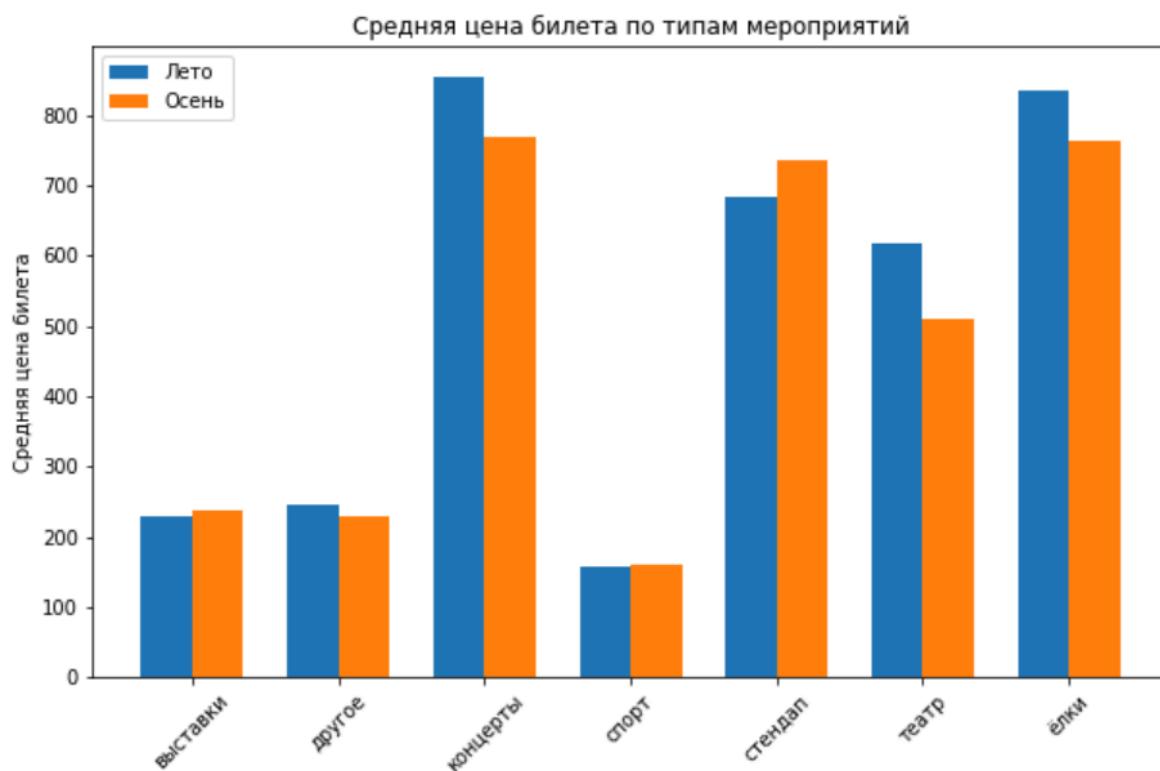
Доли заказов по возрастной категории



## 4.3 Средняя цена билета по типам мероприятий: лето vs осень

В большинстве категорий летние цены выше или близки к осенним.

Наиболее заметное снижение осенью — **театр** (примерно -15% и более), рост — **стендап**.



## 5. Осенняя активность пользователей (сентябрь–октябрь 2024)

Посчитаны дневные метрики:

- orders\_count (кол-во заказов),
- DAU (уникальные пользователи),
- orders\_per\_user (заказов на пользователя),
- avg\_ticket\_price (средняя цена билета).

**Наблюдения:**

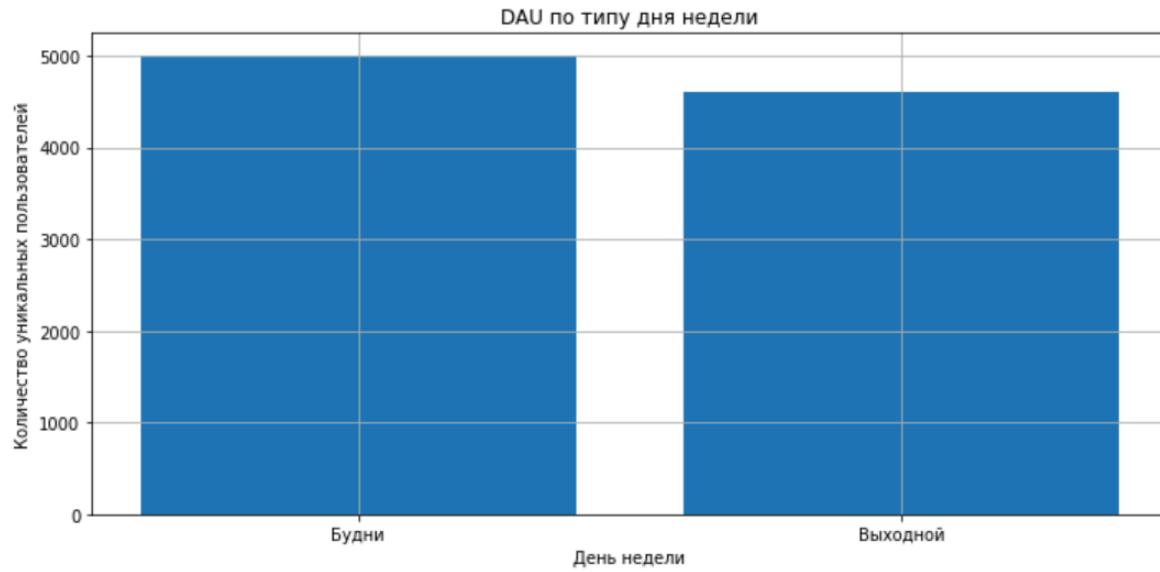
- orders\_count и DAU растут к концу октября;

- `orders_per_user` и средняя цена билета стабильны, иногда наблюдаются пики;
- пики объяснимы акциями (снижение цены → рост заказов).



## 5.1 Недельная сезонность

В будние дни активность немного выше, чем в выходные.



## 6. Популярные регионы, события и партнёры

## 6.1 Регионы

Лидеры по числу событий и заказов:

- Каменевский регион
- Североярская область  
(вместе дают значимую долю заказов)

## 6.2 Партнёры

Топ-3 партнёра по заказам/выручке:

- “Билеты без проблем”
- “Лови билет!”
- “Билеты в руки”

Они формируют большую долю заказов и выручки: зона внимания для бизнеса.

---

## 7. Проверка гипотез: mobile vs desktop (осень 2024)

### Важно про корректность сегментов

Некоторые пользователи встречаются в обоих сегментах устройств (mobile и desktop) → такие пользователи исключены из сравнения, чтобы группы были независимыми.

---

### Гипотеза 1: mobile-пользователи делают больше заказов

**H0:** среднее число заказов на пользователя одинаково (mobile = desktop)

**H1:** mobile > desktop

Использован **t-тест Уэлча** (equal\_var=False).

**Результат:** p-value ≈ 0 → различие статистически значимо.

**Вывод:** пользователи с мобильных устройств действительно делают больше заказов.

---

### Гипотеза 2: различается среднее время между заказами (days\_since\_prev)

Анализ выполнен только для пользователей с 2+ заказами (иначе days\_since\_prev отсутствует).

**Результат t-теста:** p-value = 1.0 → различий не обнаружено.

(Здесь можно добавить небольшой комментарий: распределение может быть не нормальным → лучше проверять Манна–Уитни)

---

## 8. Итоги и рекомендации бизнесу

### Итоги

- осенью растут DAU и количество заказов, при этом средняя цена билета и заказы на пользователя в целом стабильны;
- наиболее популярные категории — концерты и театр;
- цены осенью чуть ниже летних, но рост активности компенсирует выручку;
- mobile-пользователи делают больше заказов, чем desktop;
- время между заказами статистически не отличается.

### Рекомендации

- учитывать сезонный рост осенью при планировании маркетинга и нагрузок;
- усилить работу с топ-партнёрами и параллельно развивать партнёров второго эшелона;
- обратить внимание на категории: театр (сильное снижение цены) и стендап (рост цены) — возможны разные стратегии продвижения.